



2883

PATENT TRADEMARK OFFICE

103.1038.01

This application is submitted in the name of the following inventor(s):

<u>Inventor</u>	<u>Citizenship</u>	<u>Residence City and State</u>
Banga, Gaurav	India	Sunnyvale, California

The assignee is Network Appliance, Inc., a California corporation having an office at 495 East Java Drive, Sunnyvale California 94089.

Title of the Invention

Auto-Detection of Duplex Mismatch on an Ethernet

Background of the Invention

1. Field of the Invention

This invention relates to auto-detection of a communication mismatch, such as in a networking environment.

2. *Related Art*

In communication systems, it is often necessary to configure differing devices at remote ends of communication network with matching communication parameters. One common circumstance in which this can be important occurs when two devices are coupled using a LAN (local area network), such as an Ethernet, but are logically located relatively remotely. For example, the two devices might include an end-host and a switch, maybe belonging to either different organizations or different administrative domains within a single organization. A parameter mismatch may occur when the devices treat the communication link as either half-duplex or full-duplex where a first one of the devices treats the communication link as half-duplex, while a second one of the devices will treat the communication link as full-duplex. When devices are configured so that such a duplex mismatch occurs, substantial degradation in communication bandwidth and other performance characteristics often results.

One problem with known systems using Ethernet protocols is that the Ethernet protocol standard does not contain sufficient logic to auto-detect and to resolve such parameter mismatches. In consequence, attempting to determine the cause of, and attempting to correct, performance problems that originate as a result of a protocol mismatch can be quite difficult. This process is generally manual and often involves inspection of the configurations of both communicating devices. Because the communicating devices often belong to either different organizations or different administered domains

1 within a single organization, parameter mismatches, particularly duplex mismatches, can
2 occur quite often. Duplex mismatches can lead to significant loss of time on the part of
3 system administrators, loss of effective communication for a length of time, and a rela-
4 tively excessive number of calls for technical support.

5
6 Accordingly, it would be advantageous to provide a technique for auto-
7 detection of communication mismatches that is not subject to drawbacks of the known art.

8 9 Summary of the Invention

10
11 The invention provides a method and system for auto-detection of commu-
12 nication mismatches, such as in a networking environment. A device using a communi-
13 cation protocol uses a technique for protocol augmentation (similar to that described in
14 the Incorporated Disclosure) to determine sufficient information about whether there is a
15 protocol parameter mismatch (such as, for example, a duplex parameter mismatch,) and to
16 determine how to adjust its protocol parameters so that the parameter mismatch is obvi-
17 ated. In a preferred embodiment, the protocol includes an Ethernet protocol, and the
18 mismatch includes information about whether the devices at the end of a communication
19 link are using half-duplex or full-duplex settings. A first device using the Ethernet gener-
20 ates messages that force any one of a set of second devices using the same Ethernet to
21 generate responsive messages to send to the first device; the first device determines, by
22 examining features of the responsive messages from the responding set of second devices,

1 what protocol settings the set of second devices is using. With this information, the first
2 device can adjust its protocol parameter settings to match the responding second device.
3 In a preferred embodiment, the procedure is best used with a set of second devices that
4 includes five or more responding devices.

5
6 The invention provides an enabling technology for a wide variety of appli-
7 cations for computer assisted automatic error detection and diagnosis of communication
8 parameters, so as to obtain substantial advantages and capabilities that are novel and non-
9 obvious in view of the known art. Examples described below primarily relate to auto-
10 detection of duplex mismatch on an Ethernet, but the invention is broadly applicable to
11 many different types of communication and networking systems.

12 Brief Description of the Drawings

13
14
15 Figure 1 (collectively including figure 1A and figure 1B) shows a block
16 diagram of a system for auto-detection of duplex mismatch on an Ethernet.

17
18 Figure 2 shows a process flow diagram of a method for operating a system
19 for auto-detection of duplex mismatch on an Ethernet.

Detailed Description of the Preferred Embodiment

In the following description, a preferred embodiment of the invention is described with regard to preferred process steps and data structures. Embodiments of the invention can be implemented using general-purpose processors or special purpose processors operating under program control, or other circuits, adapted to particular process steps and data structures described herein. Implementation of the process steps and data structures described herein would not require undue experimentation or further invention.

Lexicography

The following terms refer or relate to aspects of the invention as described below. The descriptions of general meanings of these terms are not intended to be limiting, only illustrative.

- **error detection and diagnosis** — In general, a technique for detecting errors and other failures, and for determining a likely cause thereof
- **lower-level** and **higher-level** protocols — In general, these terms refer to a relationship between two protocols, particularly to their relationship as a higher-level protocol which relies on operation of a lower-level protocol and which is able to

1 alter parameters of the lower-level protocol, not necessarily to any particular pro-
2 tocols.

- 3
- 4 • **manipulating parameters** — In general, a technique for using a higher-level
5 protocol to determine whether a lower-level protocol is operating relatively effi-
6 ciently using a set of selected parameters for the lower-level protocol, and using
7 the lower-level protocol to repeatedly and rapidly alter those selected parameters
8 so as to find an optimal set of selected parameters.

- 9
- 10 • **monitoring statistics** — In general, information regarding performance of the file
11 server or other device.

- 12
- 13 • **network protocol** — In general, a technique for communication between devices,
14 such as, for example, between: (a) the file server or other device; and (b) a point
15 external to the file server or other device.

- 16
- 17 • **protocol augmentation** — In general, a technique for using a higher-level proto-
18 col to determine whether a lower-level protocol is operating relatively efficiently
19 using a set of selected parameters for the lower-level protocol, and using the
20 lower-level protocol to repeatedly and rapidly alter those selected parameters so as
21 to find an optimal set of selected parameters.

1 As noted above, these descriptions of general meanings of these terms are
2 not intended to be limiting, only illustrative. Other and further applications of the inven-
3 tion, including extensions of these terms and concepts, would be clear to those of ordinary
4 skill in the art after perusing this application. These other and further applications are
5 part of the scope and spirit of the invention, and would be clear to those of ordinary skill
6 in the art, without further invention or undue experimentation.

7
8 *Related Application*

9
10 This application is able to use technology disclosed in the following docu-
11 ments:

- 12
13 • U.S. Patent Application Serial No. 09/456,027, filed December 12, 1999, in the
14 name of the same inventor, titled "Computer Assisted Automatic Error Detection
15 and Diagnosis of File Servers", attorney docket number NAP-042.

16
17 This document is hereby incorporated by reference as if fully set forth
18 herein. This document is sometimes referred to herein as the "Incorporated Disclosure."

1 *System Elements*

2
3 Figure 1 (collectively including figure 1A and figure 1B) shows a block
4 diagram of a system for auto-detection of duplex mismatch on an Ethernet.
5

6 A system 100 includes a first device 110, a communication network 120,
7 and a set of second devices 130.
8

9 The first device 110 can include any device capable of communication us-
10 ing an Ethernet protocol, and capable of carrying out the procedures described herein. In
11 a preferred embodiment, the first device 110 includes a computer having a processor,
12 program and data memory, mass storage, and coupled to the communication network 120.
13 As used herein, the term “computer” is intended in its broadest sense, and includes any
14 device having a programmable processor or otherwise falling within the generalized
15 Turing machine paradigm.
16

17 The communication network 120 includes any technique for sending infor-
18 mation between the file server 110 and at least one point outside the file server 110. In a
19 preferred embodiment, the communication network 120 includes a LAN, such as an
20 Ethernet. In alternative embodiments, the communication network 120 can include an-
21 other type of computer network, such as an Internet, intranet, extranet, or a virtual private
22 network, or a non-computer network, such as a direct communication line, a switched

1 network such as a telephone network, or some combination thereof. In such alternative
2 embodiments, the communication network 120 would likely include some other commu-
3 nication protocol other than an Ethernet protocol.

4
5 Similar to the first device 110, the set of second devices 130 can include
6 any device capable of communication using an Ethernet protocol, and capable of carrying
7 out the procedures described herein. In a preferred embodiment, the set of second devices
8 130 includes computers having a processor, program and data memory, mass storage, and
9 coupled to the communication network 120. When a member of the set of second devices
10 130 responds to the first device 110, that second device is termed herein a "responding
11 second device."
12

13 Figure 1A shows a block diagram of a first use of the system 100.
14

15 In a first use of the system 100, the first device 110 sends a first message
16 111 (called herein a "reverse packet trigger" message), using the communication network
17 120, to the set of second devices 130. As described below, the reverse packet trigger
18 message 111 prompts any number of the set of second devices 130 to generate a second
19 message 131 (called herein an "induced packet" message) in response to the reverse
20 packet trigger message 111. The responding devices in the set of second devices 130 thus
21 generate and send a sequence of induced packet messages 131, using the communication
22 network 120, back to the first device 110. The first device 110 is thus able to measure a

1 response to the reverse packet trigger message 111 from the responding devices in the set
2 of second devices 130, such as by counting the number of induced packet messages 131
3 received from any responding device at the first device 110. This count is relevant to pa-
4 rameter settings at the responding second devices 130.

5
6 Figure 1B shows a block diagram of a second use of the system 100.

7
8 In a second use of the system 100, the first device 110 sends the reverse
9 packet trigger message 111, using the communication network 120, to a responding de-
10 vice in the set of second devices 130. Similar to figure 1A, the reverse packet trigger
11 message 111 prompts a responding second device 130 to generate an induced packet mes-
12 sage 131 in response to reverse packet trigger message 111. The responding second de-
13 vice 130 thus generates a sequence of induced packet messages 131, using the communi-
14 cation network 120, back to the first device 110.

15
16 In this second use of the system 100, the first device 100 sends a sequence
17 of third messages 112 (called herein "jam packet" messages), using the communication
18 network 120, to the set of responding second devices 130. If either the first device 110 or
19 a responding second device 130 is configured for half-duplex communication using the
20 communication network 120, the jam packet messages 112 will interfere with the induced
21 packet messages 131, thus reducing the number of induced packet messages 131 received
22 at the first device 110. The first device 110 is thus able to measure the response to the

reverse packet trigger message 111 from a responding second device 130, such as by counting the number of induced packet messages 131 from that responding second device received at the first device 110. The first device 110 is thus also able to determine a difference between the number of induced packet messages 131 received at the first device 110 under conditions in which jam packet messages 112 either are or are not present on the communication network 120. This difference is also relevant to parameter settings at the responding second devices 130.

Method of Operation

Figure 2 shows a process flow diagram of a method for operating a system for auto-detection of duplex mismatch on an Ethernet.

A method 200 includes a set of flow points and a set of steps. The system 100 performs the method 200. Although the method 200 is described serially, the steps of the method 200 can be performed by separate elements in conjunction or in parallel, whether asynchronously, in a pipelined manner, or otherwise. There is no particular requirement that the method 200 be performed in the same order in which this description lists the steps, except where so indicated.

A portion of the method 200 from the flow point 210 to the flow point 220 corresponds to the first step described above with regard to figure 1A.

1
2 At a flow point 210, the first device 110 is ready to determine parameter
3 settings for a number of devices included in a set of second devices 130 coupled to the
4 communication network 120. The number of devices included in the set of second de-
5 vices 130 is preferably five or greater. When a member of the set of second devices 130
6 responds to the first device 110, that second device is termed herein the "responding sec-
7 ond device."

8
9 At a step 211, the first device 110 sends the reverse packet trigger message
10 111, using the communication network 120, to the responding second device 130. As
11 part of this step, the communication network 120 attempts to deliver the reverse packet
12 trigger message 111 to the responding second device 130. As part of this step, the re-
13 sponding second device 130 attempts to receive the reverse packet trigger message 111.

14
15 In a preferred embodiment, the reverse packet trigger message 111 can in-
16 clude any packet, or sequence of packets, which when received by the responding second
17 device 130, would have the effect of causing the responding second device 130 to gener-
18 ate a message back to the first device 110 in response. For example, the reverse packet
19 trigger message 111 can include an ICMP ECHO request, a layer 2 PING message, or
20 some other message to which, according to the protocol used on the communication net-
21 work 120, the responding second device 130 must respond.

1 At a step 212, if the responding second device 130 has received the reverse
2 packet trigger message 111, the responding second device 130 generates an induced
3 packet message 131 in response to the reverse packet trigger message 111. As part of this
4 step, the communication network 120 attempts to deliver the induced packet message 131
5 to the first device 110. As part of this step, the first device 110 attempts to receive the
6 induced packet message 131.

7
8 The responding second device 130 thus generates and sends a sequence of
9 induced packet messages 131, using the communication network 120, back to the first
10 device 110. The first device 110 is thus able to measure a response to the reverse packet
11 trigger message 111 from the responding second device 130, such as by counting the
12 number of induced packet messages 131 received at the first device 110. This count is
13 relevant to parameter settings at the responding second device 130.

14
15 The first device 110 repeats the step 211 and the step 212 for a length of
16 time, sufficient to acquire information regarding a number of induced packet messages
17 131 received by the first device 110 from the responding second device 130.

18
19 At a flow point 220, the first device 110 is thus able to determine a number
20 of induced packet messages 131 sent by the responding second device 130 in response to
21 the sequence of reverse packet trigger messages 111 sent by the first device 110.

1 A portion of the method 200 from the flow point 220 to the flow point 230
2 corresponds to the second step described above with regard to figure 1B.

3
4 At a step 221, the first device 110 sends the reverse packet trigger message
5 111, using the communication network 120, to the responding second device 130. As
6 part of this step, the communication network 120 attempts to deliver the reverse packet
7 trigger message 111 to the responding second device 130. As part of this step, the re-
8 sponding second device 130 attempts to receive the reverse packet trigger message 111.

9
10 At a step 222, the first device 110 also sends the jam packet message 112,
11 using the communication network 120, to the responding second device 130. As part of
12 this step, the communication network 120 attempts to deliver the jam packet message 112
13 to the responding second device 130. As part of this step, the responding second device
14 130 attempts to receive the jam packet message 112.

15
16 At a step 223, if the responding second device 130 has received the reverse
17 packet trigger message 111, the responding second device 130 generates an induced
18 packet message 131 in response to the reverse packet trigger message 111. As part of this
19 step, the communication network 120 attempts to deliver the induced packet message 131
20 to the first device 110. As part of this step, the first device 110 attempts to receive the
21 induced packet message 131.

1 The responding second device 130 thus generates and sends a sequence of
2 induced packet messages 131, using the communication network 120, back to the first
3 device 110. However, if the responding second device 130 is configured to treat the
4 communication network 120 as half-duplex, jam packet messages 112 present on the
5 communication network 120 cause the responding second device 130 to delay sending
6 induced packet messages 131 until the jam packet messages 112 are no longer present.

7
8 The first device 110 repeats the step 212, the step 222, and the step 223 for
9 a length of time, sufficient to acquire information regarding a number of induced packet
10 messages 131 received by the first device 110 while jam packet messages 112 are present
11 on the communication network 120.

12
13 At a flow point 230, the first device 110 is able to measure a response to the
14 reverse packet trigger message 111 while jam packet messages 112 are present on the
15 communication network 120, such as by counting the number of induced packet messages
16 131 received at the first device 110. This count is relevant to parameter settings at the re-
17 sponding second device 130.

18
19 At a step 231, the first device 110 uses the measures from the flow point
20 220 and the flow point 230 to determine protocol parameters used by the responding sec-
21 ond device 130 relating to half-duplex or full-duplex use of the communication network
22 120. This step includes the following sub-steps:

- 1
- 2 • At a sub-step 231(a), the first device 110 determines how it has configured its own
- 3 protocol parameters for the communication network 120. These protocol parame-
- 4 ters can be either half-duplex or full-duplex.
- 5
- 6 • At a sub-step 231(b), the first device 110 determines whether there is a relatively
- 7 normal number of induced packet messages 131 received from the responding sec-
- 8 ond device 130.
- 9
- 10 • At a sub-step 231(c), the first device 110 determines whether there is a relatively
- 11 large number of collisions between induced packet messages 131 received from
- 12 the responding second device 130 and jam packets 112 sent by the first device 110.
- 13 As part of this sub-step, the first device 110 determines whether a substantial per-
- 14 centage of these collisions are late collisions.
- 15
- 16 • If the first device 110 has configured its own protocol parameters as half-duplex,
- 17 then the responding second device 130 will have a protocol mismatch only if the
- 18 responding second device 130 has configured its own protocol parameters as full-
- 19 duplex. In this case, the first device 110 will see a relatively large number of colli-
- 20 sions as indicated in sub-step 231(c). If the first device 110 has configured its own
- 21 protocol parameters as full-duplex, then the responding second device 130 will
- 22 have a protocol mismatch only if the second device 130 has configured its own

1 protocol parameters as half-duplex. In this case, the first device 110 will see a
2 relatively small number of induced packet messages 131.

3
4 At a flow point 240, the first device 110 has thus determined protocol pa-
5 rameters used by the responding second device 130 relating to half-duplex or full-duplex
6 use of the communication network 120, and whether those protocol parameters match
7 corresponding protocol parameters used by the first device 110.

8
9 At a step 241, the first device 110 repeats the steps from the flow point 210
10 through and including the flow point 240 a number of times, so that any traffic anomalies
11 on the communication network 120 are accounted for. In a preferred embodiment, the
12 first device 110 repeats those steps about three times, each time determining whether or
13 not there is a protocol mismatch, and adjusting its protocol parameters (as described be-
14 low with regard to step 242) in response to a majority vote of results.

15
16 At a step 242, the first device 110 adjusts its protocol parameters to match
17 corresponding protocol parameters used by the responding second device 130. In alter-
18 native embodiments, the first device 110 may cause an operator to adjust protocol pa-
19 rameters used by the responding second device 130 so as to match corresponding protocol
20 parameters used by the first device 110.

1 At a flow point 250, the first device 110 and the responding second device
2 130 are thus using matching protocol parameters relating to half-duplex or full-duplex use
3 of the communication network 120.

4
5 The method 200 is performed one or more times starting from the flow
6 point 210 and continuing therefrom. In a preferred embodiment, the first device 110 re-
7 peatedly performs the method 200, starting from the flow point 210 and continuing there-
8 from, so as to periodically and continuously determine that there is no parameter mis-
9 match between the first device 110 and the responding second device 130. If the set of
10 responding second devices 130 is fewer than five in number, the method 200 might be
11 performed a greater number of times to determine statistically relevant results. However,
12 for the invention to provide its advantages, there is no particular requirement for such
13 repetition, and the method 200 need only be performed on initial connectivity between the
14 first device 110 and the responding second device 130.

15 16 *Generality of the Invention*

17
18 The invention has general applicability to various fields of use, not neces-
19 sarily related to the techniques described above. For example, these fields of use can in-
20 clude automatic error detection and diagnosis of communication parameters for other
21 types of devices, other communication links, and other communication protocols.

1 Other and further applications of the invention in its most general form, will
2 be clear to those skilled in the art after perusal of this application, and are within the
3 scope and spirit of the invention.

4
5 *Alternative Embodiments*

6
7 Although preferred embodiments are disclosed herein, many variations are
8 possible which remain within the concept, scope, and spirit of the invention, and these
9 variations would become clear to those skilled in the art after perusal of this application.